Multidimensional Assessment of India's Manufacturing GVC Linkages

Chapter 1 of *Learning & Upgrading in GVCs: An Analysis of India's Manufacturing Sector*

Prologue…

- Most of the production process are composed of different tasks performed by several geographically separated economic agents with its profound impact similar to what observed regarding the division of labour.

- This system of global production sharing, which is known as Global Value Chains (GVCs), is eventually forming an extensive network of economic values or value added that explains the changing nature of international trade & development.

- To analyse India's position, functions, specialisation & value addition of manufacturing GVCs, it is required to quantify the extent, drivers, and impacts of India's Manufacturing links in GVCs.

- This overall broad objective can be transformed into three fundamental questions:
  - What is the Extent of India's Manufacturing Links in GVCs?
  - What are the Determinants of India's Manufacturing Links in GVCs?
  - What are the Impacts of India's Manufacturing Links in GVCs?

What is the Extent of India's Manufacturing Links in GVCs?

- GVC Participation Using Gross Trade Data & Trade in Value Added (TiVA) Data.

- Analysis of value-added trade requires much more sophisticated data like TiVA and WIOD.

- In this context, the gross nature of imports and exports data is such that it cannot provide any direct information about the foreign and domestic value added to an economy.

- Nonetheless, we can get some indirect indications about the GVC phenomenon through a set of analytic refinements based on gross trade data and its informed classifications.

GVC Participation Using Gross Trade Data

- The whole purpose of this part of my research is to quantify the degree of value transformation taking place domestically between imports and exports due to India's GVC integration in the manufacturing sector.

- The consideration of using gross imports & exports data to draw a quantitative inference about India's GVC potential leads to a different conceptualisation of the existing production process regarding knowledge, learning, and upgrading.

- Traditional trade theories assume that the whole production process of a product is taken place in one country and marketed in another. However, the notion of GVC trade is different – quantifying how much of India's export value is contributed by imports and how much India is adding value to exports. The basic concept is "import to export" or I2E.

The question I am trying to address…

- First, we have to look at what India imports and exports. Looking at the top imported and exported products of India, classified at the most disaggregated level (Harmonised System [HS] 6-digits) is a good starting point.

- Second, we have to compare statistically those product-level imports with exports regarding values, volumes, and unit prices.
    - If exports and imports follow a similar distribution or having a short distance between two distributions, this may imply that relatively little transformation is taking place in the domestic segment of India's manufacturing GVCs.

Why we should care about this question?

- The logic behind these objectives is that we can easily point out the core factor of global inequality, i.e. an asymmetric distribution of knowledge. In fact, due to Paul Romer's paper on endogenous technological change in 1990, economists finally figured out that knowledge, not capital, not technology by itself, not labour or land, was the ultimate factor of production that allowed economic growth.

- Many authors emphasise the importance of understanding knowledge (local & global) creation and utilisation in the process of innovation for value addition. They indeed identified information as the medium to move knowledge from one place to another, and its interpretation (decoding) depends upon the degree of user's (individual or organisation) set of capabilities. That is why there is a high chance of uneven distribution of knowledge in a few pockets of the world economy.

- Therefore, understanding the structure & dynamics of knowledge distribution is crucial in the age of global production networks. In this context, the processes of learning and upgrading (innovation) are the fundamental forces of diffusing productive knowledge as a means of generating economic values among the different economic actors in the various production activities.

What is my contribution?

- Collective Characterisation & Quantification

- Structure of value distribution (System's Composition)
  - The aim of this study is to discover law-governed relations between the products forming this production sharing system. In the process of this research we will be identifying the structures peculiar to this sharing system.

- Dynamics of value distribution (System's Functioning)
  - But structure is not enough to make a system. This production sharing system consists of something more than structure: it is a structure with certain properties and stability (including their movement in time, their sequence etc).

- When a structure is understood from the standpoint of its properties, it is understood as a system.

Structure of Value Distribution

How value is distributed among imports and exports?

Conceptual Framework…

- Economic development is the fruit of shared (collective) learning in an economy.
- This concept consists of two properties of knowledge and learning:
  - a limited capacity to learn individually
  - a limited capacity to transmit experiential knowledge
- These constraints are the principal drivers of team formation among individuals to unite their knowledge.
- It is the basis of collective learning, where each production process is a combination of numerous complementary capabilities.
- Economies that are participating in various production networks accumulate adequate capabilities and can combine with productivity.

Scrabble Theory of Economic Development

- The idea of limited learnability has the inference of why sometimes knowledge does not move freely and concentrate in few pockets.

- Therefore, it is hard for individuals to learn, collectives (firms, industries, countries) on the other hand can learn by gathering people.

- Hence, the diffusion of knowledge does not move necessarily because of ideas flowing into people heads, but because of people with ideas assemble in certain places.

- Capabilities embodied in the brain of people will move where they can combine with other capabilities productively.

- This conceptualisation thus indicates why knowledge accumulates slowly, and why its distribution is uneven across spaces.

Scrabble Theory of Economic Development

- The learning perspective posits that the knowledge content of a country cannot be found at the intensive margin.

- knowledge grows not by accumulating more of the same, but by adding new and different elements to existing capabilities.

- It is this evolutionary, combinatorial process that drives many economic phenomena.

Value, Information, Knowledge... Com. System

Most relevant buyers ● NAICS product 334112 ● Other

Most relevant suppliers ● NAICS product 334112 ● Other
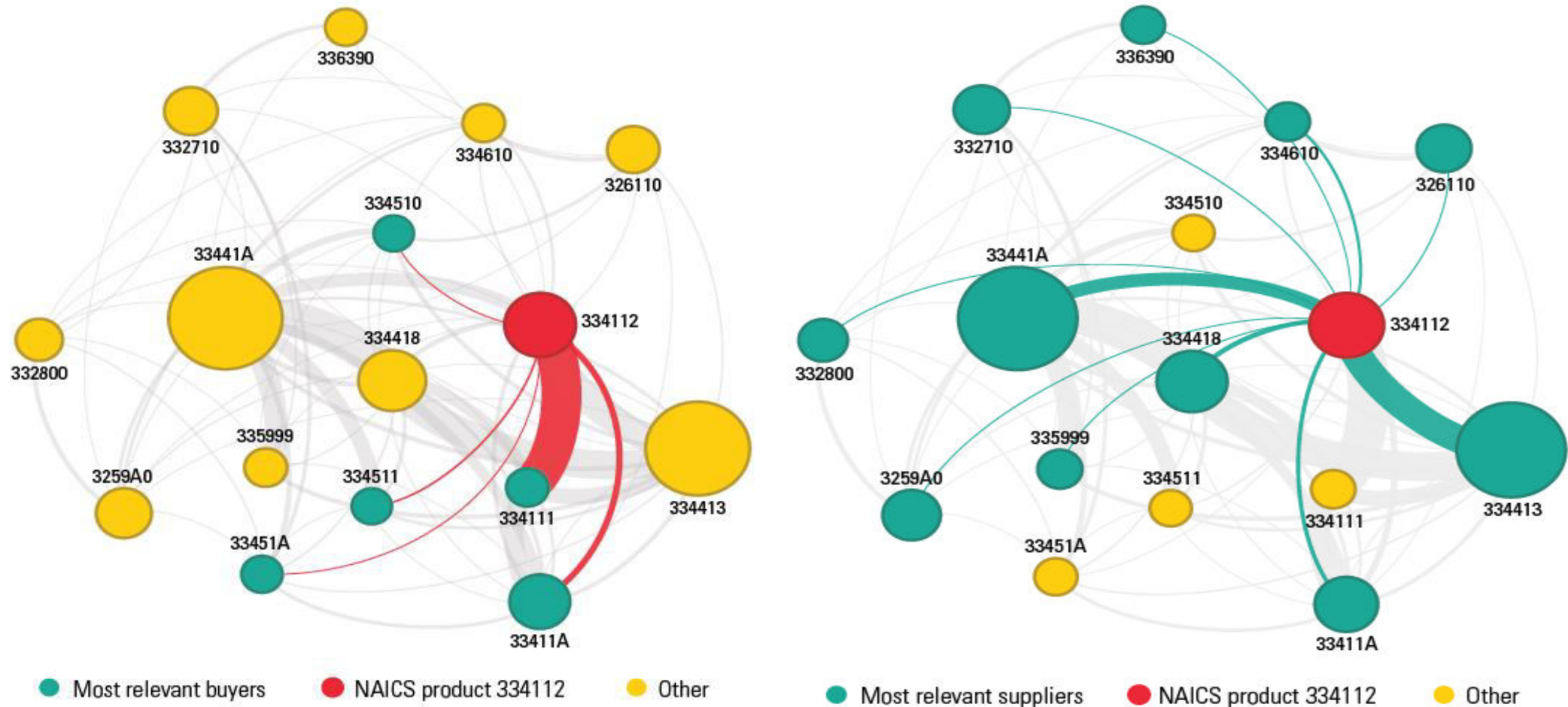
Source: Santoni and Taglioni forthcoming. Adapted from Benchmark Input-Output Data, Bureau of Economic Analysis, U.S. Department of Commerce.

Top Imports & Exports 2017

Rank	ProductCode	ProductDescription	Gross Imports USD	Share %
1	270900	Petroleum oils and oils obtained fr	60185357589	17.92248604
2	710812	Gold in unwrought forms non-monetar	26261310581	7.820306984
3	710231	Diamonds non-industrial unworked or	14782330311	4.402002734
4	270119	Other coal, not agglomerated, nes	14338312734	4.269779563
5	851790	Parts of electrical apparatus for l	8824138549	2.627723857
6	710239	Diamonds non-industrial nes excludi	7067977422	2.104759891
7	852520	Transmission apparatus, for radiote	5954849455	1.773283578
8	271111	Natural gas, liquefied	5445032100	1.621466013
9	151110	Crude palm oil	3533865436	1.052343236
10	880240	Aircraft nes of an unladen weight e	3445142934	1.025922726

| Rank	ProductCode	ProductDescription	Gross Exports USD	Share %
1	271000	Petroleum oils, etc, (excl. crude);	24149427743	11.13729458
2	710239	Diamonds non-industrial nes excludi	16904178148	7.795911924
3	300490	Other medicaments of mixed or unmix	7301144309	3.367160324
4	711319	Art. of jewellery and pts thereof o	6813783663	3.1423981
5	100630	Semi-milled or wholly milled rice	5163992817	2.381543357
6	30613	Frozen shrimps and prawns	3862418478	1.781279989
7	711311	Art. of jewellery and pts thereof o	3029016185	1.396929397
8	20230	Frozen boneless bovine meat	2993946715	1.38075597
9	870322	Automobiles with reciprocating pist	2544538865	1.173496913
10	760110	Aluminium unwrought, not alloyed	2028157679	0.935350923

Asymmetric Distribution in Top Imports & Exports

Gabaix, X. (1999). Zipf's law for cities: an explanation. The Quarterly journal of economics, 114(3), 739-767.
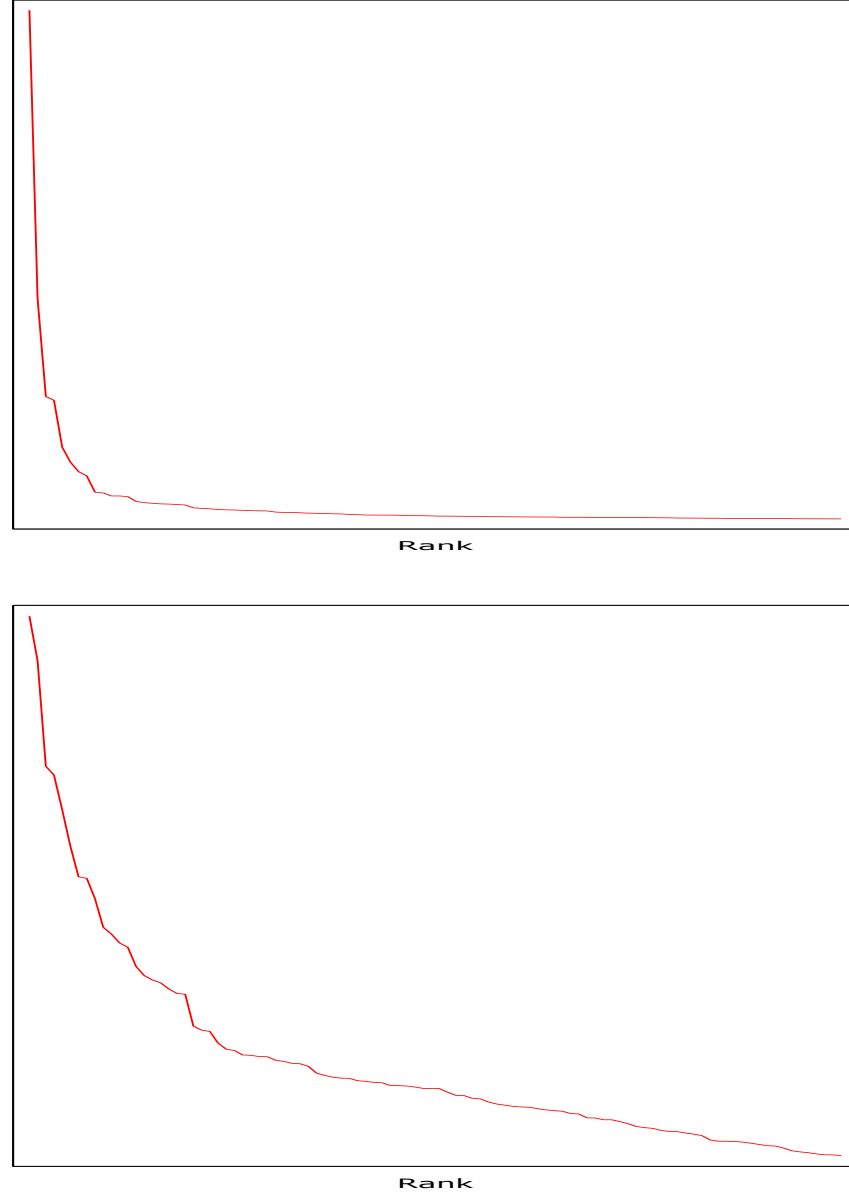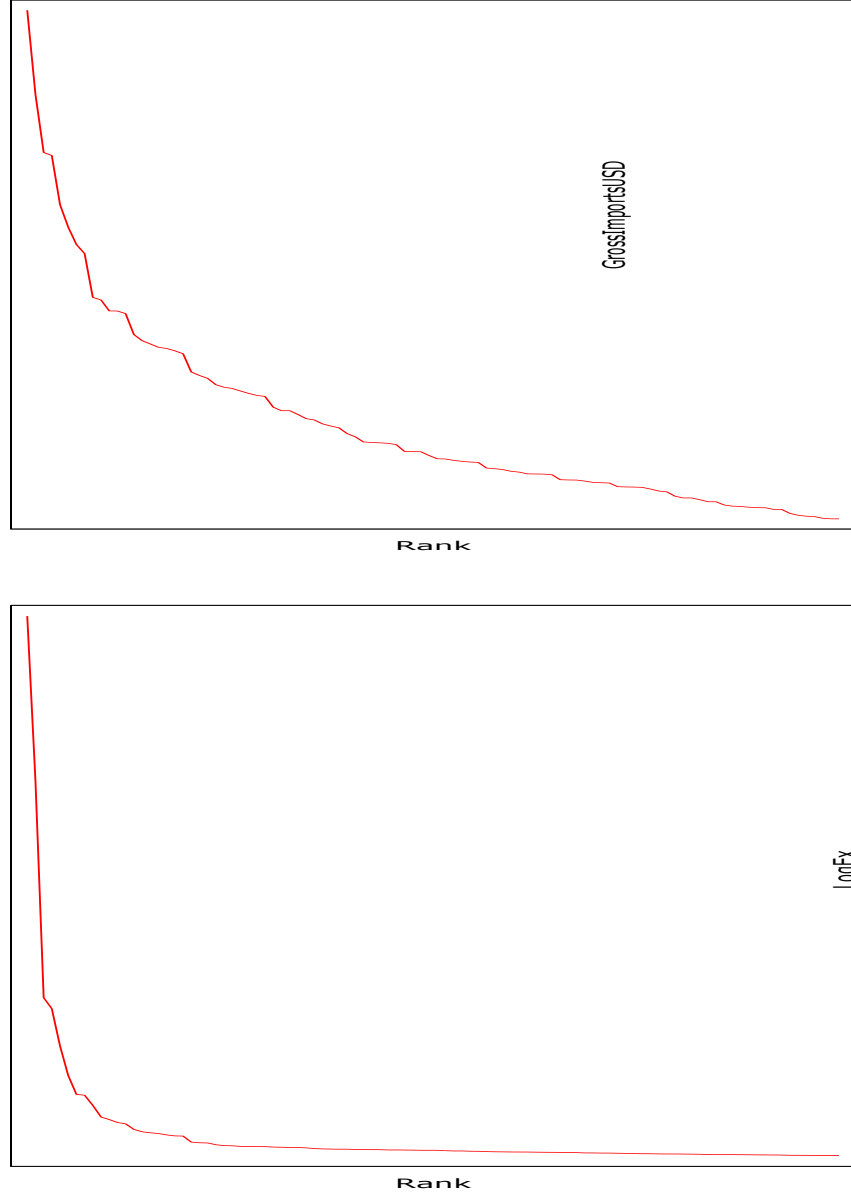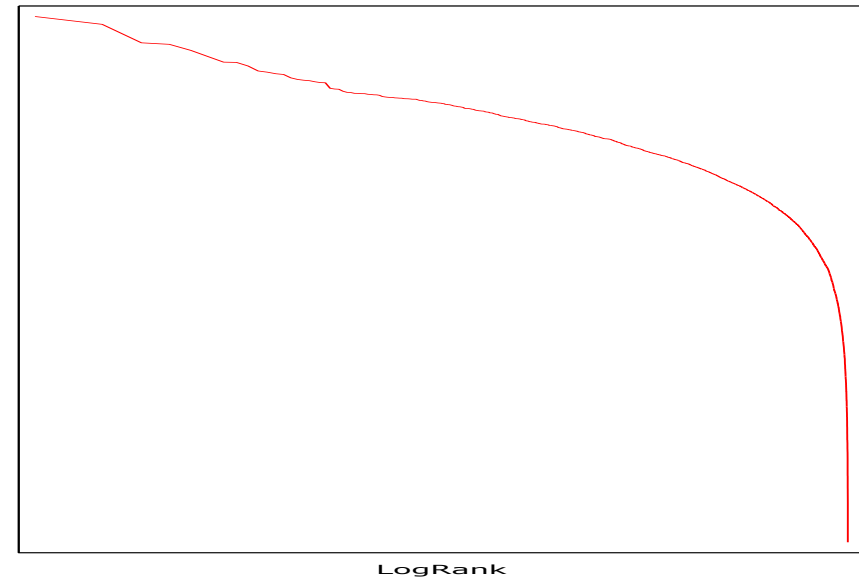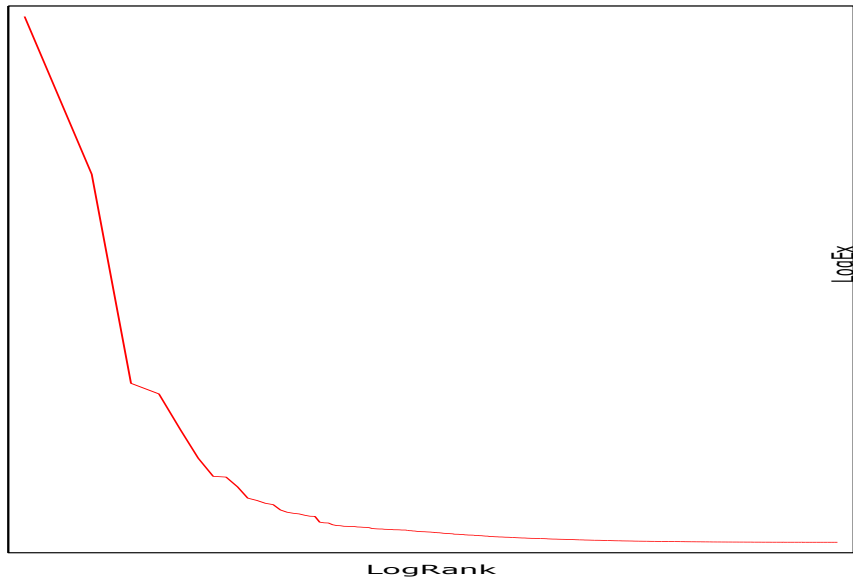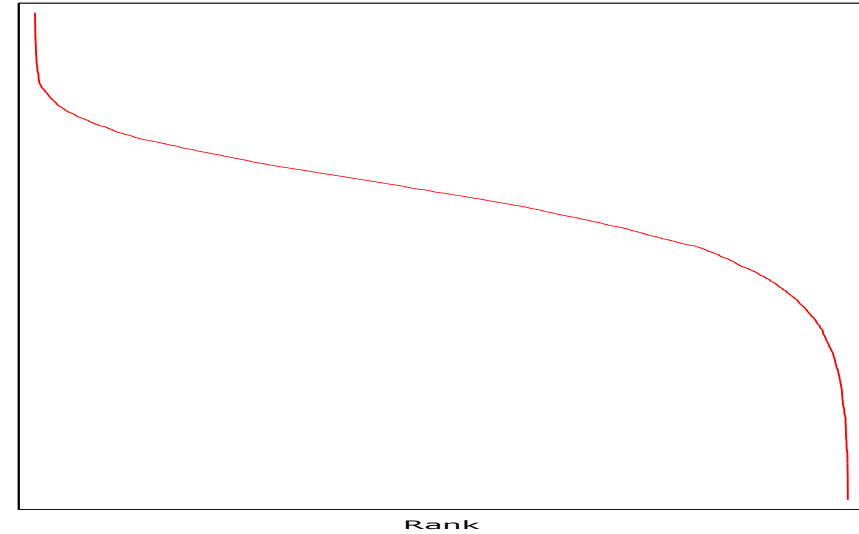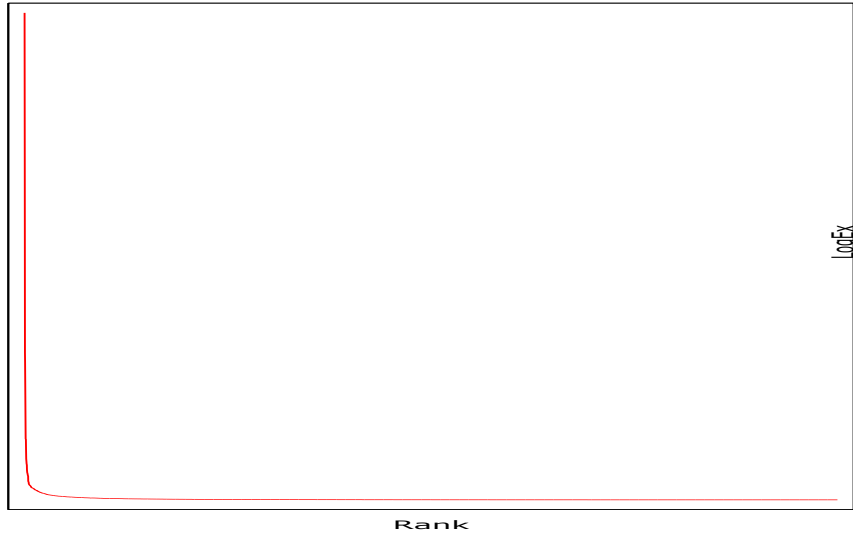
Gabaix, X. (1999). Zipf's Law and the Growth of Cities. American Economic Review, 89(2), 129-132.

Rank-Size Distribution of All Imports
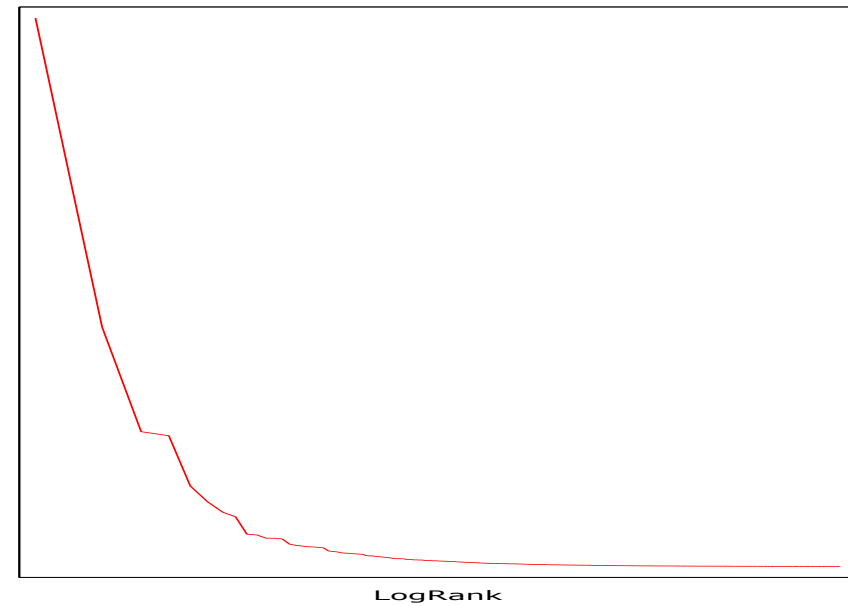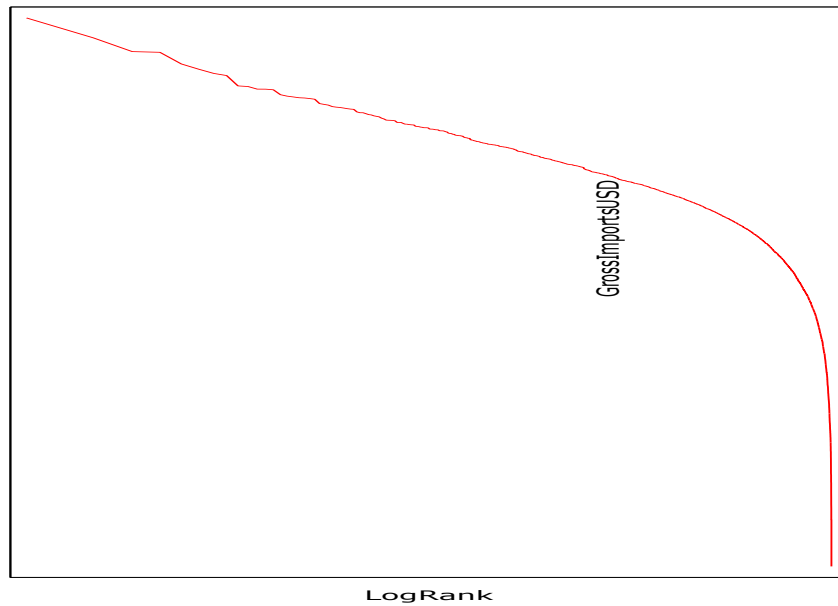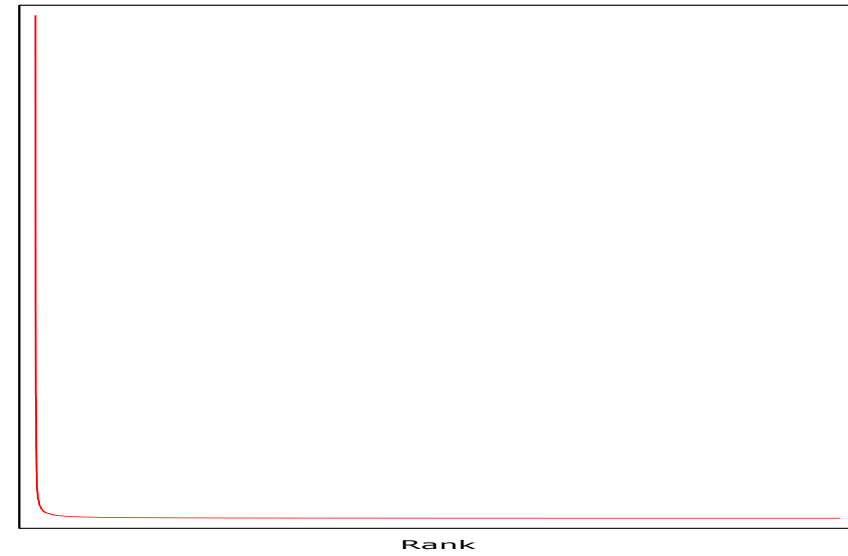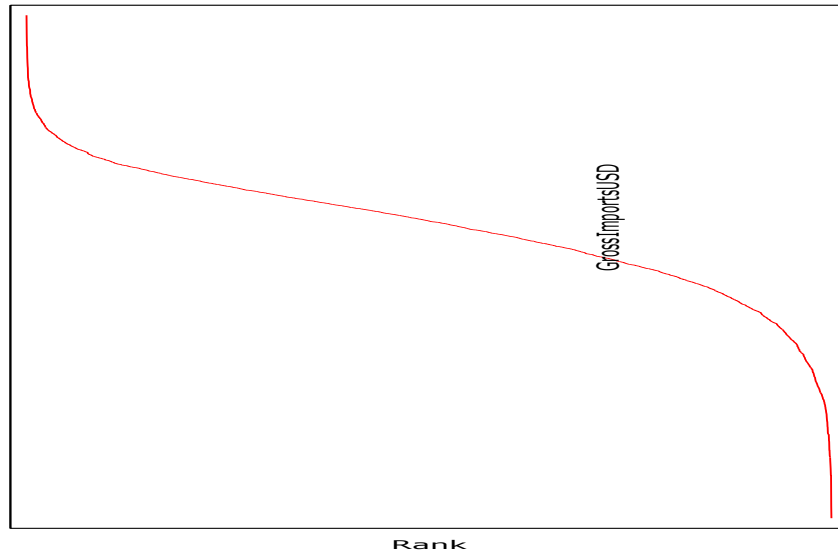
Rank-Size Distribution of All Exports

Asymmetric Value Distribution

- A country's distribution of exports tends to follow a lognormal, power, or Zipf 's law distribution. Zipf 's law, originally applied to language, states that given some universe of items, the frequency of any item is inversely proportional to its rank in the frequency table.

- That is, a few items account for the bulk of the given universe; the contribution of most items is marginal. Exports and imports loosely follow such asymmetric distribution laws.

- Therefore, the marginal additional information that can be gathered from import and export products beyond the top 50 to 100 is generally small.

Power Laws… Asymmetric Value Distribution

- $P(x) = Cx^{-\alpha} = \dfrac{C}{x^\alpha}$, for $x \geq x_{\min}$

- Normalisation $(\alpha > 1)$

$$1 = \int_{x_{\min}}^{\infty} P(x)\, dx = C \int_{x_{\min}}^{\infty} \frac{dx}{x^\alpha} = \frac{C}{\alpha - 1}\, x_{\min}^{-\alpha+1} \Rightarrow c = (\alpha - 1)x_{\min}^{\alpha-1}$$

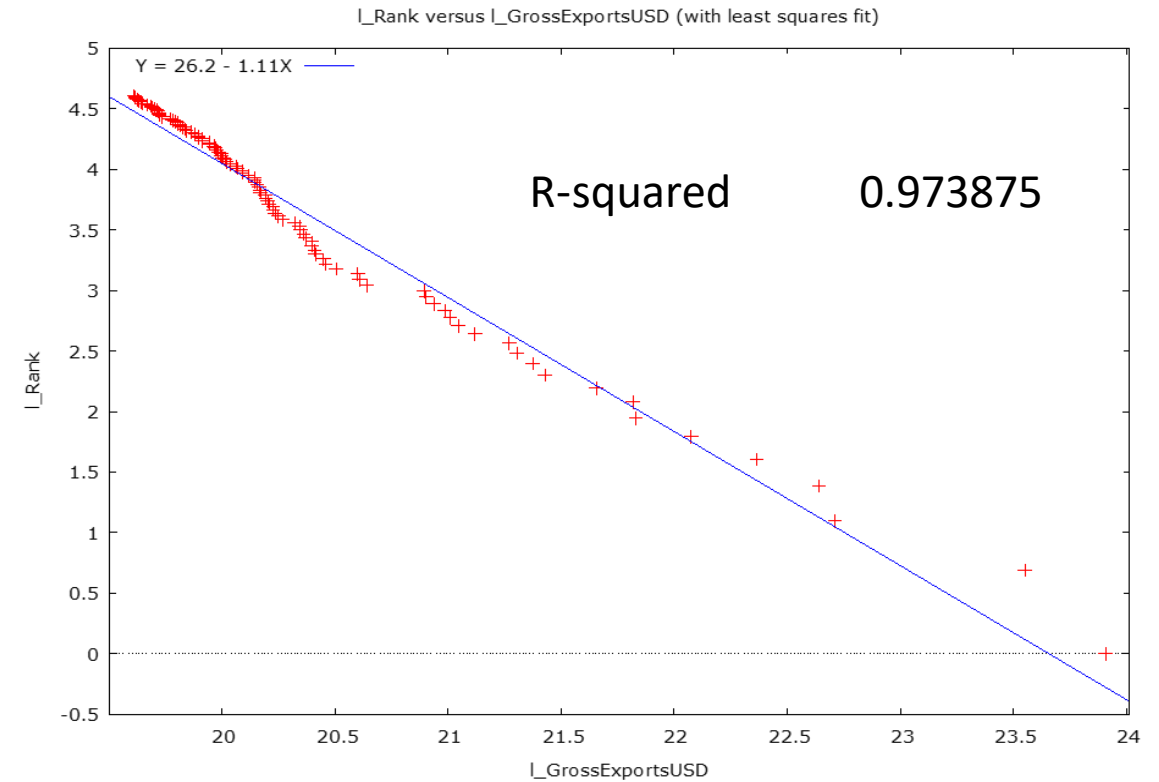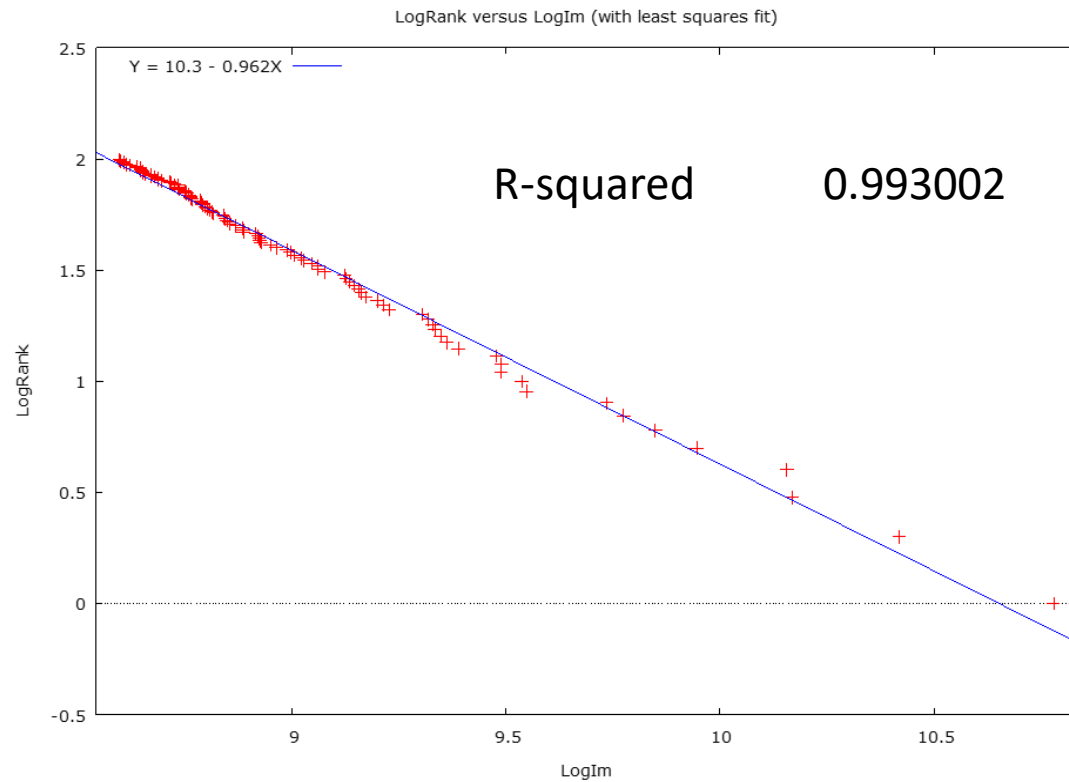- Power law PDF: $P(x) = \dfrac{\alpha - 1}{x_{\min}} \left(\dfrac{x}{x_{\min}}\right)^{-\alpha}$

- Complementary CDF: $\bar{F}(x) = Pr(X > x) = \int_{x}^{\infty} P(x)\, dx = \left(\dfrac{x}{x_{\min}}\right)^{-(\alpha-1)}$

$$\bar{F}(x) = \bar{C}x^{-(\alpha-1)} \Rightarrow \log \bar{F}(x) = \log \bar{c} - (\alpha - 1)\log x$$

Some Papers on Power Laws

- Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, *30*(1), 185-206.

- Gabaix, X. (2009). Power laws in economics and finance. *Annu. Rev. Econ.*, *1*(1), 255-294.

- Di Giovanni, J., Levchenko, A. A., & Ranciere, R. (2011). Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics*, *85*(1), 42-52.

- Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics letters*, *74*(1), 15-19.

- Di Giovanni, J., & Levchenko, A. A. (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, *120*(6), 1083-1132.
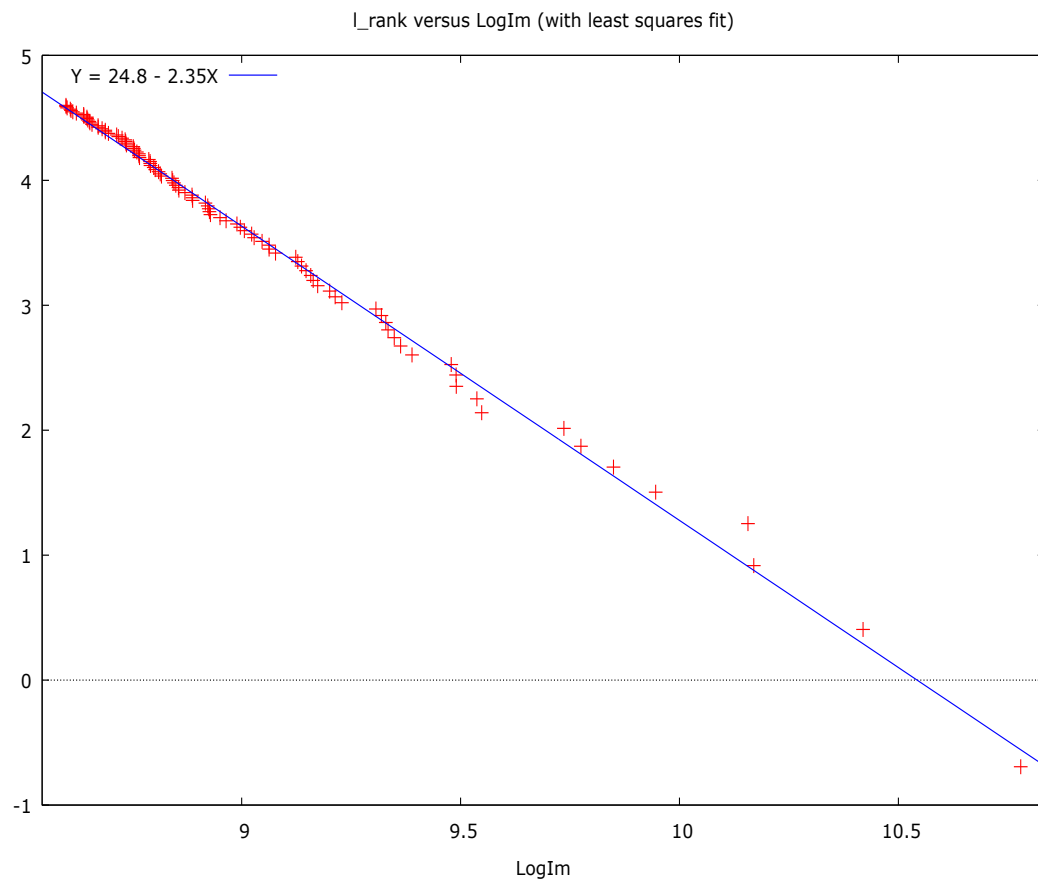
Scale Invariance

- Scaling of the density: $x \rightarrow bx, P(bx) = c(bx)^{-\alpha} = b^{-\alpha}cx^{-\alpha} \propto P(x)$

- Scale invariance: $\dfrac{P(100x)}{P(10x)} = \dfrac{P(10x)}{P(x)}$

LogRank versus LogIm (with least squares fit)

Y = 10.3 - 0.962X

R-squared   0.993002

l_Rank versus l_GrossExportsUSD (with least squares fit)

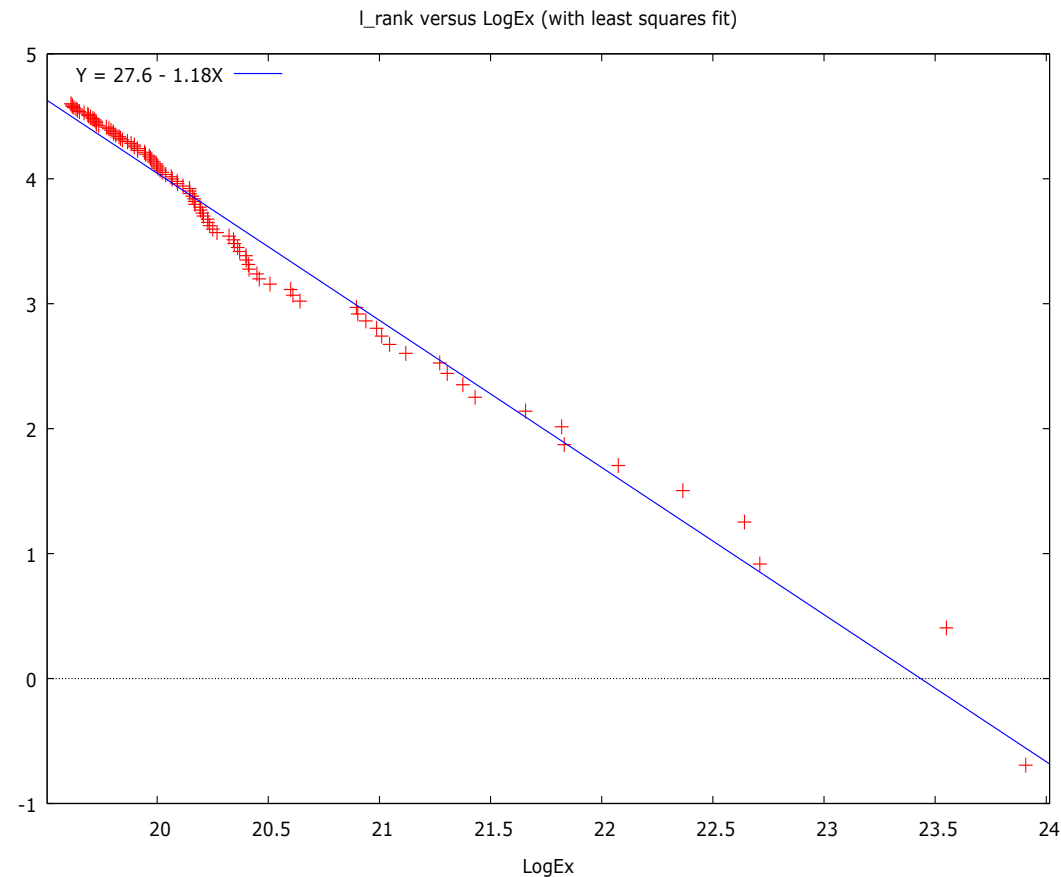Y = 26.2 - 1.11X

R-squared   0.973875

Robust Estimation

Gabaix, X., & Ibragimov, R. (2011). Rank– 1/2: a simple way to improve the OLS estimation of tail exponents. Journal of Business & Economic Statistics, 29(1), 24-39.

l_rank versus LogIm (with least squares fit)

$Y = 24.8 - 2.35X$

l_rank versus LogEx (with least squares fit)

$Y = 27.6 - 1.18X$

R-squared   0.995994

R-squared   0.980881

Some Quotes

- "Few if any economists seem to have realized the possibilities that such invariants hold for the future of our science. In particular, nobody seems to have realized that the hunt for, and the interpretation of, invariants of this type might lay the foundations for an entirely novel type of theory"… Schumpeter (1949, p. 155), about the Pareto law

- "More empirically, the other area that I regard as promising but is not being pushed very much, it is not very popular, is looking at distributions of things: distributions of price changes, distributions of size of firms, distributions of size of cities and so on and so forth. There's a small intellectual community that's interested in that and the striking thing is, of course, that most of these distributions tend to have very fat tails. Now this reflects an underlying heterogeneity in the population, and what's wrong with macroeconomics in my opinion is that the populations are modeled as if there were a representative consumer. The representative consumer model masks a great deal of problems, and I think these two are connected: distributions have fatter tails than we expect from models, and models are based on the assumption of a representative consumer. The fat tails mean that a relatively small number of events, or people or something, have a big influence. And I think most of our theories of price changes, of changes in investment in response to different conditions are deficient because they don't take account of the shapes of the distributions."… Kenneth Arrow (2005)

High Optimisation & Preferential Attachment ("rich-gets-richer" or "fit-gets-richer")

$P(x)$ - Distribution of values

$A(x)$ - Value size (information)

$C(x)$ - Cost scale as $A^\alpha(x)$, α sets the relative weight of products of different size.

The expected cost: $$E(A^\alpha) = \int_X p(\mathbf{x}) A^\alpha(\mathbf{x}) d\mathbf{x}.$$

$R(x)$ - Resource which restricts the sizes of products

Limitation on the total quantity of the resources: $$\int_X R(\mathbf{x}) d\mathbf{x} = \kappa,$$

Function & Dynamics of Value Distribution

What is the implication of power law?

Concentration of Value

- Let X_i be the export value of product i. This variable follows a power law or Pareto distribution if the complementary cumulative distribution function is: $P(X_i > x) = \dfrac{k}{x^\alpha}, \alpha > 0$

- Pareto distribution only applies for values equal to or greater than some minimum $x_{\min} > 0, k = (x_{\min})^\alpha$

- The exponent α, also known as the Pareto or tail exponent, is the key parameter. It provides a single measure of how heavy the tails are, i.e. how concentrated the variable is in the top units.

- The lower the exponent α is, the fatter the tails are, and the greater the inequality existing between the top units or the greater their concentration. Zipf's law states that a is approximately equal to one.
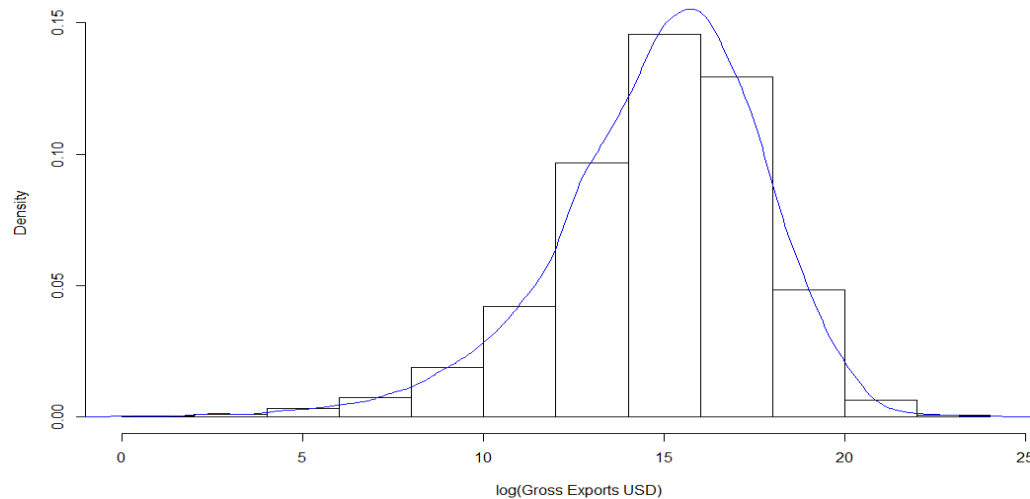
Non-parametric Distribution

Summary Statistics of Log (Gross Exports USD 2017)
(4199 valid observations)

Mean	Median	Minimum	Maximum
14.838	15.162	0.00000	23.908
Std. Dev.	C.V.	Skewness	Ex. kurtosis
2.8642	0.19304	-0.75966	1.2022
5% Perc.	95% Perc.	IQ range	Missing obs.
9.5807	18.931	3.5989	0

Summary Statistics of Log (Gross Imports USD 2017)
(4199 valid observations)

Mean	Median	Minimum	Maximum
15.178	15.509	1.9459	23.991
Std. Dev.	C.V.	Skewness	Ex. kurtosis
2.6996	0.17787	-0.65079	0.91548
5% Perc.	95% Perc.	IQ range	Missing obs.
10.273	18.983	3.4684	0

Exports 2017

Imports 2017
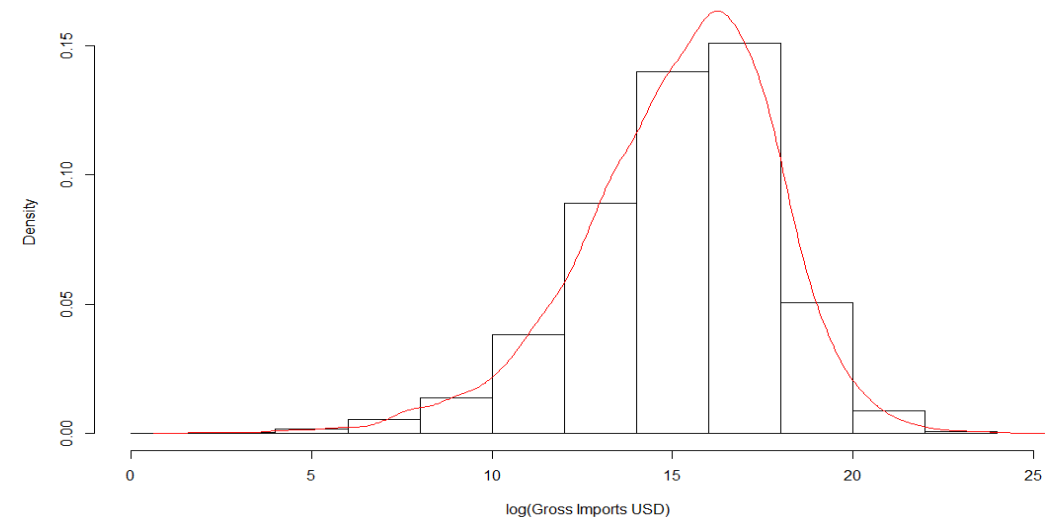
Generalized beta distribution of the second kind
$GB2(a, b, p, q)$

▶ Density:

$$f(x; a, b, p, q) = \frac{a}{bB(p, q)} \frac{(x/b)^{ap-1}}{(1 + (x/b)^a)^{p+q}},$$

where

- ▶ $B(p, q)$ is the beta function,
- ▶ $b > 0$ is a scale parameter,
- ▶ $p > 0, q > 0$ and $a > 0$ are shape parameters.

- McDonald, J. (1984): Some Generalized Functions for the Size Distribution of Income. Econometrica, 52 (3), pp. 647-663.
- Jenkins, S. P. (2009). Distributionally-sensitive inequality indices and the GB2 income distribution. *Review of Income and Wealth, 55*(2), 392-398.

- Value quintile share ratio (QSR)
  - QSR is the ratio of the sum of the upper quintile values over the sum of the lower quintile values.
- Gini index (GINI)
  - The index is an inequality indicator measuring the expected absolute difference between two independently selected values relative to the mean value.

Cumulative distribution plots: GB2 vs empirical d.f.

Density plots: Kernel density estimate (Epanechnikov) and GB2 density.

Epanechnikov kernel:
- ▶ Quadratic weight function within an interval around each observed value.
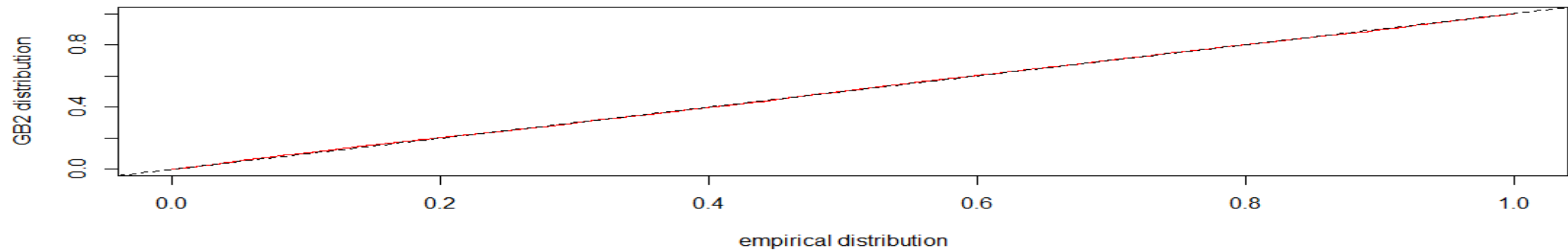- ▶ Length of interval = band width.
- ▶ N = sample size

GB2 Distribution of Exports 2017

Cumulative Distribution plot

Density plot

N = 4198   Bandwidth = 0.4551
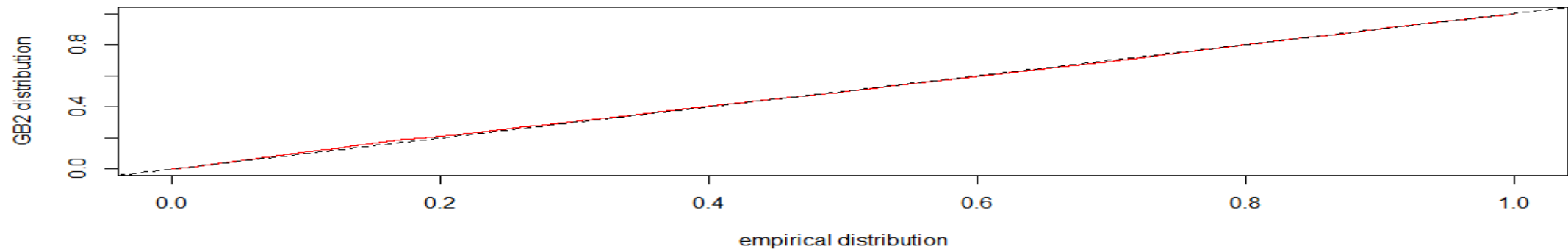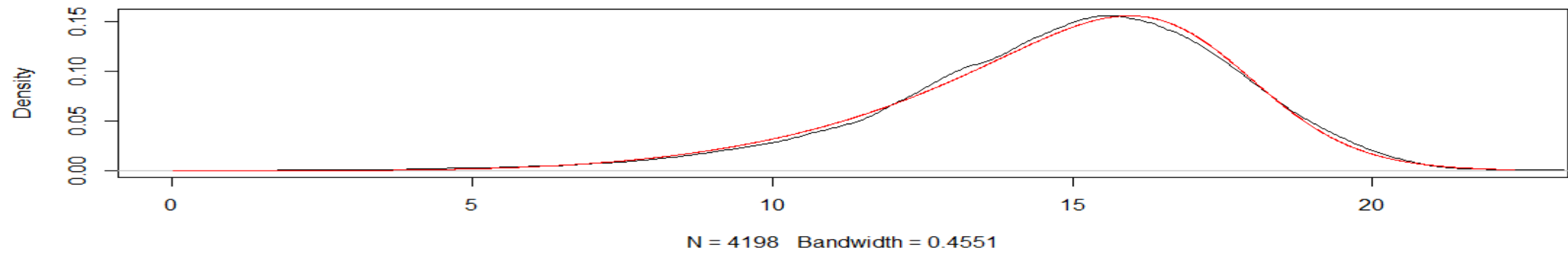
Value as Information

- Conceptually, exported or imported **value** can be thought of as **information** being stored in or transmitted as exported or imported products that can take on different values. A **export or import** can be thought of as a unit of storage that can take on, at different times, one of several different specified product values, following some process for taking on those product values.

- Informally, we get information from a product by looking at its value, just as we get information from an email by reading its contents. In the case of the product, the information is about the process behind the product.

Entropy

- The **entropy** of export or import is the "**amount of information**" contained in the products. This amount is determined not just by the number of different product values the export or import can take on, just as the information in an email is quantified not just by the number of words in the email or the different possible words in the language of the email. Informally, the amount of information in an email is proportional to the amount of "surprise" its reading causes.

- For example, if an email is simply a repeat of an earlier email, then it is not informative at all. On the other hand, if say the email reveals the outcome of a research, then it is highly informative. Similarly, the information in export or import is tied to the amount of surprise that value of the export or import causes when revealed.

Statistical distance

- **Entropy H(X)**

$$H(X) = -\sum_{i=1}^{n} P(x_i) * log_b(P(x_i))$$

- **Conditional-Entropy H(X | Y)**

$$H(Y|X) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) * log_b(P(x_i)/P(x_i, y_j))$$

- **Mutual Information I(X,Y)**

$$MI(X,Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) * log_b(P(x_i, y_j)/(P(x_i) * P(y_j)))$$

- **Kullback-Leibler Divergence**

$$KL(P||Q) = \sum_{i=1}^{n} P(p_i) * log_2(P(p_i)/P(q_i)) = H(P,Q) - H(P)$$